

Pendekatan Interpretable Machine Learning untuk Analisis Keberhasilan Kampanye Pemasaran Menggunakan CatBoost dan SHAP

Aprilisa Arum Sari^{1*}, Oktalia Kumala Sari²

¹Program Studi Teknik Informatika, Universitas Duta Bangsa

²Program Studi Magister Ilmu Komputer, Universitas Budi Luhur

*aprilisa_arumsari@udb.ac.id

Abstrak

Memprediksi keberhasilan kampanye pemasaran digital merupakan tantangan penting akibat kompleksitas interaksi berbagai variabel, seperti alokasi anggaran dan pemilihan saluran iklan. Penelitian ini bertujuan mengembangkan model analitik pemasaran yang memiliki tingkat akurasi tinggi sekaligus mampu memberikan interpretasi yang jelas terhadap hasil prediksi. Penelitian menggunakan dataset *SalesMind Marketing Campaigns 2026* yang mensimulasikan 3.478 data kampanye pemasaran digital tahun 2026. Dataset terdiri dari variabel kategorikal seperti *ad_channel* dan *campaign_type*, serta variabel numerik seperti *marketing_spend*, *impressions*, dan *conversion_rate* sebagai target prediksi. Pendekatan yang digunakan adalah *Interpretable Machine Learning* dengan mengombinasikan algoritma CatBoost untuk memprediksi tingkat konversi (*conversion rate*) serta metode SHAP (*SHapley Additive exPlanations*) untuk menganalisis kontribusi masing-masing variabel. Proses optimasi model dilakukan menggunakan GridSearchCV, menghasilkan performa yang sangat baik dengan nilai RMSE sebesar 0,0012, MAE sebesar 0,005, serta koefisien determinasi (R^2) sebesar 99,12%. Hasil analisis menunjukkan bahwa alokasi anggaran merupakan faktor paling dominan dalam meningkatkan konversi, tanpa menunjukkan indikasi penurunan efektivitas marginal. Selain itu, penggunaan platform interaktif seperti Meta dan TikTok terbukti berperan dalam meningkatkan efektivitas kampanye. Temuan penelitian ini memberikan kontribusi dalam menyediakan model prediktif yang akurat sekaligus informatif, sehingga dapat mendukung pengambilan keputusan strategis dalam pengelolaan pemasaran digital secara lebih efektif.

Kata kunci : Pemasaran Digital, Machine Learning, CatBoost, SHAP, Interpretable AI, CRISP-DM

Abstract

Predicting the success of digital marketing campaigns remains a significant challenge due to the complex interactions among various variables, such as budget allocation and advertising channel selection. This study aims to develop a marketing analytics model that achieves high predictive accuracy while also providing clear interpretability of the prediction results. The study uses the SalesMind Marketing Campaigns 2026 dataset, which simulates 3,478 digital marketing campaign records from 2026. The dataset consists of categorical variables such as ad_channel and campaign_type, as well as numerical variables including marketing_spend, impressions, and conversion_rate as the prediction target. The proposed approach applies Interpretable Machine Learning by combining the CatBoost algorithm to predict conversion rates and SHAP (SHapley Additive exPlanations) to analyze the contribution of each variable. Model optimization was performed using GridSearchCV, resulting in excellent performance with an RMSE of 0.0012, an MAE of 0.005, and a coefficient of determination (R^2) of 99.12%. The analysis results indicate that budget allocation is the most dominant factor in improving conversion rates without showing indications of diminishing marginal effectiveness. In addition, the use of interactive platforms such as Meta and TikTok significantly contributes to campaign effectiveness. These findings contribute to providing an accurate and informative predictive model that can support strategic decision-making in digital marketing management more effectively.

Keywords : Digital Marketing, Machine Learning, CatBoost, SHAP, Interpretable AI, CRISP-DM

1. Pendahuluan

Di era ekonomi digital saat ini, lanskap bisnis mengalami transformasi besar yang digerakkan oleh ketersediaan mahadata (*big data*). Skala investasi perusahaan pada sektor pemasaran digital secara global terus melonjak drastis secara tahun ke tahun (*year-on-year*)^[1]. Menurut proyeksi pertumbuhan industri, total pengeluaran iklan digital dunia yang tercatat mencapai kisaran \$640 miliar pada akhir tahun 2025 diestimasikan akan terus berekspansi melampaui angka \$700 miliar di penghujung tahun 2026^{[2][3]}. Namun, terlepas dari injeksi kapital yang masif pada periode 2025–2026 tersebut, tingkat efisiensi kampanye sering kali masih jauh dari ekspektasi^[4]. Berbagai laporan industri mencatat bahwa rata-rata tingkat konversi (*conversion rate*) global lintas sektor sering kali stagnan di rentang yang sangat rendah, yakni hanya berkisar antara 2% hingga 5%^[5]. Ketimpangan antara besarnya anggaran yang dikeluarkan dengan rendahnya persentase konversi yang dihasilkan ini menjadikan kemampuan untuk mengukur dan memprediksi tingkat pengembalian investasi (ROI) sebagai elemen yang sangat krusial bagi keberlangsungan operasi bisnis^{[6],[7],[8]}. Tingkat konversi tidak lagi sekadar menjadi indikator keberhasilan, melainkan kompas strategis yang memandu perusahaan dalam mendistribusikan sumber daya mereka di masa mendatang^{[8][9]}. Namun, metode analisis statistik tradisional

seperti regresi linier maupun logistik sering kali kesulitan untuk memodelkan dinamika pasar modern. Perilaku konsumen dan respons terhadap iklan tidak lagi bergerak dalam garis lurus, melainkan membentuk pola hubungan non-linier yang sangat kompleks antara pengeluaran anggaran, jenis kampanye yang diluncurkan, dan hasil konversi akhir^{[10][11]}. Menghadapi keterbatasan ini, penerapan algoritma *machine learning* tingkat lanjut, khususnya yang berbasis pada ansambel pohon keputusan (*decision tree ensembles*) seperti *gradient boosting*, mulai diadopsi secara luas^{[12],[13]}. Berbagai studi membuktikan bahwa metode komputasi canggih ini mampu memberikan tingkat akurasi prediksi yang jauh lebih superior dalam menangkap pola-pola tersembunyi dari data historis pemasaran^[14]. Untuk menjembatani kesenjangan antara akurasi komputasi dan kemudahan interpretasi strategis tersebut, penelitian ini mengusulkan sebuah pendekatan *Interpretable Machine Learning*. Secara spesifik, penelitian ini menerapkan algoritma *Categorical Boosting* (CatBoost) sebuah algoritma yang secara inheren tangguh dalam memproses variasi data kategorikal pemasaran secara otomatis dan menggabungkannya dengan kerangka kerja interpretasi SHAP (*SHapley Additive exPlanations*).

Tujuan utama dari penelitian ini adalah merancang arsitektur analitik prediktif yang tidak

hanya mampu memperkirakan tingkat konversi pemasaran dengan margin kesalahan yang minim, tetapi juga secara transparan membongkar alasan matematis di balik tebakan tersebut. Melalui pendekatan ini, diharapkan kalkulasi komputasi yang rumit dapat diterjemahkan menjadi wawasan bisnis yang terukur, rasional, dan mudah dipahami oleh para pengambil keputusan.

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Penelitian terkait mengenai analisis faktor penentu keberhasilan kampanye pemasaran telah banyak dilakukan dengan berbagai pendekatan, khususnya menggunakan metode machine learning dan data mining. Berbagai penelitian menunjukkan bahwa algoritma berbasis boosting seperti *XGBoost* dan *CatBoost* memiliki kemampuan yang tinggi dalam meningkatkan akurasi prediksi dibandingkan metode statistik tradisional.

1. Penelitian oleh Yujia Yang (2025) dengan judul *A Study on Bank Marketing Prediction Based on XGBoost* menunjukkan bahwa algoritma *XGBoost* memiliki performa yang sangat baik dalam memprediksi keberhasilan pemasaran perbankan. Hasil penelitian menunjukkan bahwa model *XGBoost* mampu mengungguli beberapa algoritma klasifikasi tradisional dengan nilai akurasi sebesar 83,39%, AUC sebesar 0,9162,

dan recall sebesar 0,8247. Selain itu, *XGBoost* dinilai efektif dalam menangani data berdimensi tinggi dan hubungan nonlinier pada data pemasaran pelanggan bank^[15].

2. Temuan serupa juga disampaikan oleh Jun Cui (2023) dalam penelitian *Digital Marketing Program Design Based on Abnormal Consumer Behavior Data Classification and Improved Homomorphic Encryption Algorithm* yang menunjukkan bahwa penerapan algoritma klasifikasi data perilaku konsumen mampu meningkatkan efektivitas strategi pemasaran digital. Penelitian tersebut menekankan bahwa analisis perilaku konsumen berbasis data dapat membantu perusahaan dalam mengidentifikasi pola interaksi pelanggan, meningkatkan akurasi pengambilan keputusan pemasaran, serta mendukung optimalisasi program pemasaran digital secara lebih adaptif dan efisien. Selain itu, penggunaan pendekatan algoritma yang ditingkatkan juga dinilai mampu menjaga keamanan data konsumen tanpa mengurangi performa analisis data pemasaran^[16].

3. Kemampuan *CatBoost* dalam menangani data kompleks juga ditunjukkan pada penelitian Yu (2025) yang menghasilkan akurasi sebesar 90,91% dan nilai AUC sebesar 93,82% dalam prediksi keberhasilan pelanggan produk perbankan^[17]. Sementara itu, Faiaz et al. (2025) berhasil mencapai akurasi sebesar 99,85% dalam prediksi persetujuan pinjaman dengan

memanfaatkan kombinasi *CatBoost* dan SHAP untuk mengidentifikasi fitur dominan seperti *risk score* dan *debt-to-income ratio*[18]. Penelitian lain oleh Okoye et al. (2025) menunjukkan bahwa integrasi *Explainable Artificial Intelligence* dalam analisis perilaku konsumen mampu menghasilkan model dengan akurasi 89,2% dan nilai AUC 0,91, di mana faktor seperti durasi sesi, kedalaman pencarian, dan riwayat pembelian terbukti berpengaruh signifikan terhadap konversi pelanggan[19].

4. Selain itu, Liu (2026) mengembangkan model *hybrid ensemble learning* berbasis SHAP yang mampu meningkatkan akurasi prediksi sekaligus memberikan interpretasi yang lebih baik terhadap faktor-faktor yang memengaruhi keputusan pembelian konsumen[20]. Penelitian oleh Cheng (2026) menunjukkan bahwa kombinasi *CatBoost* dan SHAP mampu menghasilkan performa prediksi yang tinggi sekaligus mengungkap faktor-faktor utama yang memengaruhi konsentrasi sedimen tersuspensi. Hasil analisis SHAP berhasil menjelaskan kontribusi setiap variabel secara lebih transparan sehingga membantu memahami mekanisme yang memengaruhi hasil prediksi model[21]. Selain pendekatan *machine learning*, metode data mining klasik juga banyak diterapkan dalam analisis perilaku konsumen dan strategi pemasaran. Oktalia (2023) menggunakan algoritma Apriori untuk menemukan pola pembelian produk sembako dan menghasilkan

aturan asosiasi dengan nilai *support* dan *confidence* yang tinggi[22].

5. Penelitian oleh Aditya Wadanur (2022) juga menunjukkan bahwa algoritma Apriori dan FP-Growth mampu menghasilkan pola transaksi penjualan yang dapat digunakan sebagai dasar pengambilan keputusan strategi penjualan[23]. Di sisi lain, penelitian oleh Aprilisa (2024) membahas pengembangan aplikasi marketplace berbasis *Design Thinking* untuk mendukung digitalisasi UMKM, di mana pendekatan berbasis pengguna (*user-centered*) terbukti meningkatkan efektivitas sistem dan interaksi pengguna dalam mendukung pemasaran digital[24].

2.2. Landasan Teori

1. Machine Learning

Machine Learning adalah sub-bidang dari kecerdasan buatan yang memberikan kemampuan pada sistem komputer untuk mengidentifikasi pola data dan menyusun prediksi secara otomatis[25]. Di sektor pemasaran digital, teknologi ini berperan krusial dalam memproses dataset skala besar guna menemukan keterkaitan antarvariabel yang rumit. Implementasinya mencakup analisis mendalam terhadap perilaku pelanggan dan proyeksi angka konversi, yang pada akhirnya meningkatkan akurasi pengambilan keputusan strategis[26].

2. Interpretable Machine Learning

Interpretable Machine Learning (IML) merupakan

paradigma yang menyelaraskan akurasi prediksi dengan transparansi model. Mengingat banyak algoritma modern bersifat *black box* yang sulit dipahami secara intuitif, IML hadir untuk memberikan penjelasan logis atas output yang dihasilkan^[18]. Dengan mengidentifikasi variabel-variabel kunci yang memengaruhi hasil, pendekatan ini meningkatkan akuntabilitas dan kredibilitas data dalam mendukung proses pengambilan keputusan strategis^[27].

3. CatBoost

CatBoost (*Categorical Boosting*) merupakan algoritma berbasis *gradient boosting* dari Yandex yang dioptimalkan untuk mengolah data kategorikal secara langsung tanpa menuntut prapemrosesan yang rumit^[28]. Keunggulan utama algoritma ini terletak pada kemampuannya meminimalisir *overfitting* serta meningkatkan stabilitas model saat menghadapi dataset kompleks^[29].

4. SHAP (SHapley Additive exPlanations)

SHAP (*SHapley Additive exPlanations*) merupakan kerangka kerja interpretasi model yang mengadopsi prinsip teori permainan melalui pendekatan Shapley Value^{[8],[30]}. Metode ini menguantifikasi kontribusi setiap fitur terhadap hasil prediksi, baik dalam cakupan global maupun lokal. Secara global, SHAP mengidentifikasi variabel yang paling berpengaruh terhadap model secara menyeluruh, sedangkan secara lokal, ia merincikan logika di balik prediksi pada observasi

data tertentu. Implementasi SHAP sangat krusial untuk memperkuat transparansi model dan membantu pengguna memvalidasi akuntabilitas dari setiap keputusan yang dihasilkan^[31].

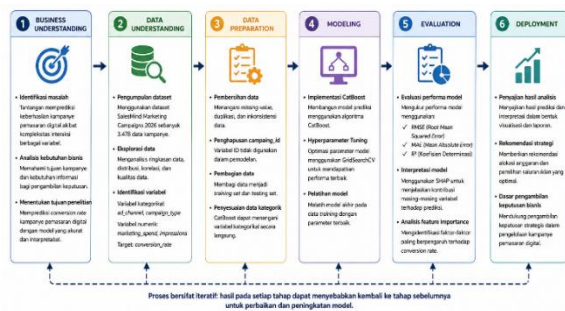
5. Pemasaran Digital

Pemasaran digital merupakan aktivitas pemasaran yang memanfaatkan media dan teknologi digital untuk mempromosikan produk atau layanan kepada konsumen. Platform digital seperti Google, Meta, TikTok, dan Email digunakan untuk menjangkau target pasar secara lebih luas dan efektif^[32]. Dalam praktiknya, keberhasilan pemasaran digital diukur menggunakan berbagai indikator, salah satunya adalah conversion rate. Pemanfaatan analisis data dan machine learning dalam pemasaran digital memungkinkan perusahaan untuk memahami perilaku konsumen, mengoptimalkan strategi promosi, serta meningkatkan efektivitas kampanye pemasaran^[33].

3. Metode Penelitian

3.1. Desain dan Tahapan Penelitian

Penelitian ini disusun secara sistematis berdasarkan kerangka kerja standar industri *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Kerangka kerja ini memastikan bahwa seluruh tahapan pengolahan data sejalan dengan tujuan utama bisnis.



Gambar 1. Kerangka Penelitian

1. Pemahaman Bisnis (*Business Understanding*)

Masalah utama dalam analitik pemasaran saat ini adalah sulitnya menentukan faktor mana yang paling dominan dalam mendorong kesuksesan sebuah kampanye akibat terbatasnya transparansi model prediksi. Oleh karena itu, penelitian ini bertujuan untuk membangun sebuah sistem yang tidak hanya akurat dalam memprediksi tingkat konversi, tetapi juga mampu memberikan penjelasan yang transparan melalui pendekatan Interpretable Machine Learning. Penjelasan matematis yang transparan ini sangat dibutuhkan oleh tim pemasaran untuk menyusun strategi anggaran yang lebih efisien dan mencegah pemborosan dana iklan.

2. Pemahaman Data (*Data Understanding*)

Penelitian ini menggunakan dataset yang mensimulasikan riwayat kampanye pemasaran digital pada tahun 2026 (*SalesMind_Marketing_Campaigns_2026.csv*). Berdasarkan eksplorasi awal, dataset ini memiliki beberapa variabel independen, baik yang berbentuk kategori seperti saluran iklan (*ad_channel*) dan tipe kampanye

(*campaign_type*), maupun yang berbentuk angka seperti besaran anggaran (*marketing_spend*) dan jumlah tayangan iklan (*impressions*). Target variabel yang ingin diprediksi dari data ini adalah persentase keberhasilan tingkat konversi (*conversion_rate*).

3. Persiapan Data (*Data Preparation*)

Pada tahap ini, data dibersihkan dan disesuaikan agar siap dimodelkan. Kolom *campaign_id* dihapus dari data karena hanya berfungsi sebagai nomor identitas unik yang tidak memiliki nilai prediksi. Selain itu, karena algoritma CatBoost mampu memproses data berbasis kategori secara langsung (*native*), proses transformasi data seperti *One-Hot Encoding* tidak perlu dilakukan. Hal ini sangat menguntungkan karena dapat mencegah pembengkakan jumlah kolom data (*curse of dimensionality*). Selanjutnya, data dibagi secara acak menjadi data latih (80%) dan data uji (20%) menggunakan benih angka acak (*random seed = 42*) untuk memastikan hasil penelitian ini dapat diulang dan divalidasi di kemudian hari.

4. Pemodelan (*Modeling*)

Proses pembuatan algoritma prediksi menggunakan modul *CatBoostRegressor*. Untuk memastikan model bekerja secara optimal dan tidak hanya menghafal data (*overfitting*), penelitian ini melakukan optimasi parameter menggunakan teknik *GridSearchCV*. Teknik ini menguji berbagai kombinasi pengaturan model

melalui validasi silang (*3-fold Cross Validation*), khususnya pada pengaturan kedalaman pohon komputasi (*depth*), kecepatan model dalam belajar (*learning_rate*), dan batas maksimal iterasi pembentukan model (*iterations*).

5. Evaluasi (Evaluation)

Kinerja model dievaluasi dari dua sisi: seberapa akurat tebakannya dan seberapa logis penjelasan di balik tebakan tersebut. Tingkat akurasi pada data uji diukur menggunakan metrik *Root Mean Squared Error* (RMSE) dan *Mean Absolute Error* (MAE) untuk melihat rata-rata kesalahan tebakan model. Koefisien determinasi (R^2) juga digunakan untuk melihat seberapa besar variasi tingkat konversi yang berhasil dipelajari oleh model. Untuk evaluasi penjelasan, metode *TreeExplainer* dari modul SHAP digunakan untuk menganalisis kontribusi setiap variabel secara visual. Gambar yang dihasilkan kemudian dianalisis untuk memastikan bahwa logika pengambilan keputusan model sejalan dengan teori dan praktik pemasaran yang masuk akal.

6. Deployment

Tahap deployment dilakukan dengan menyajikan hasil prediksi dan interpretasi model sebagai rekomendasi strategis dalam pengelolaan kampanye pemasaran digital. Hasil analisis digunakan untuk membantu pengambilan keputusan terkait alokasi anggaran, pemilihan saluran pemasaran, serta evaluasi efektivitas kampanye berdasarkan visualisasi SHAP. Selain

itu, model yang dihasilkan dapat dikembangkan lebih lanjut menjadi sistem dashboard analitik pemasaran berbasis machine learning.

4. Hasil dan Pembahasan

4.1. Hasil Penelitian

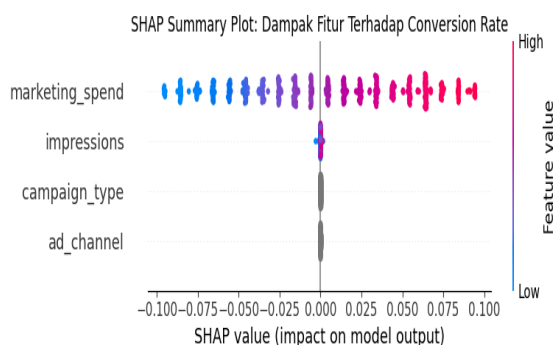
Dataset "SalesMind Marketing Campaigns 2026" yang diteliti merangkum informasi dari 3478 kegiatan kampanye digital. Berikut adalah 5 data teratas dari dataset yang digunakan:

Tabel 1. Dataset Kampanye Digital

campaign_id	ad_channel	campaign_type	marketing_spend	conversion_rate	impressions
1	Google	Brand Awareness	14907.88	0.17	49540
2	Email	Seasonal	12732.15	0.15	30439
3	Google	Flash Sale	10770.45	0.13	7338
4	TikTok	Brand Awareness	12589.94	0.15	13601
5	TikTok	Seasonal	18850.48	0.21	25260

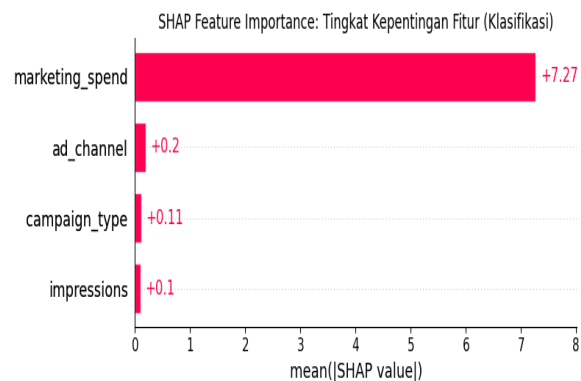
Berdasarkan ringkasan data, kampanye disebar secara cukup merata pada empat saluran pemasaran utama yaitu Google, Meta, TikTok, dan Email. Secara finansial, anggaran kampanye (*marketing_spend*) sangat bervariasi, dimulai dari kampanye skala kecil sebesar \$1.000 hingga skala besar yang mendekati angka \$20.000. Jumlah paparan iklan (*impressions*) yang dicapai juga fluktuatif, dari puluhan ribu hingga hampir 50.000 tayangan. Sementara itu, tingkat konversi (*conversion_rate*) yang dihasilkan berkisar antara 3% untuk kampanye dengan performa terendah, hingga menembus 22% pada kampanye yang

sangat sukses. Berdasarkan hasil eksperimen optimasi hiperparameter melalui *GridSearchCV*, performa optimal dari pemodelan *CatBoostRegressor* tercapai pada konfigurasi komputasional dengan kedalaman pohon (*depth*) sebesar 6, laju pembelajaran (*learning rate*) 0,05, dan batas iterasi komputasi sebanyak 300 siklus. Hasil penghitungan metrik evaluasi menunjukkan akurasi yang sangat memuaskan. Model regresi ini mencatatkan tingkat kesalahan (RMSE) sebesar 0.0012 dan MAE sebesar 0.005. Angka kesalahan yang nyaris mendekati nol ini membuktikan bahwa selisih antara tebakan model dengan persentase konversi di dunia nyata sangatlah kecil. Lebih dari itu, model ini meraih skor R^2 sebesar 0.9912. Hal ini menegaskan bahwa model mampu memahami dan memetakan 99,12. Untuk menguraikan bagaimana model bisa menghasilkan tebakan seakurat itu, nilai SHAP diekstraksi dan divisualisasikan.



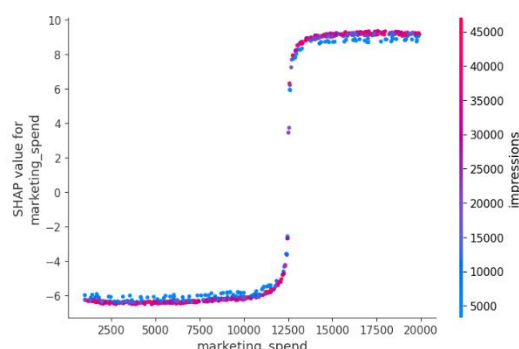
Gambar 2. Hasil Keluaran Analisis CatBoost dan SHAP Pada gambar 1 di atas menampilkan rangkuman eksekusi analisis di lingkungan Google Colab, yang menegaskan keberhasilan pelatihan model

dan generasi visualisasi SHAP. Untuk pembedahan yang lebih rinci, *Feature Importance Plot* pada Gambar 2 menyoroti derajat kontribusi masing-masing atribut.



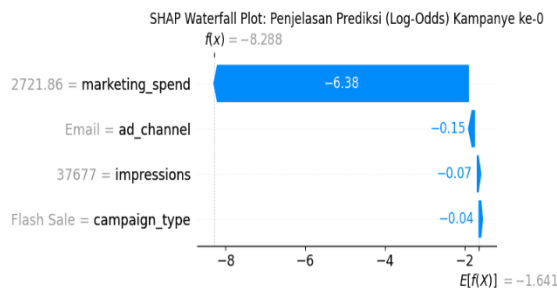
Gambar 3. Grafik SHAP Feature Importance

Pada Gambar 2, anggaran pemasaran (*marketing_spend*) teridentifikasi secara absolut sebagai determinan prediktif terkuat, jauh mengungguli atribut lainnya seperti saluran iklan (*ad_channel*) maupun tipe kampanye (*campaign_type*). Hal ini memvalidasi dominasi variabel finansial dalam menggerakkan keberhasilan metrik konversi.



Gambar 4. Grafik Visualisasi Interpretasi SHAP Gambar 3 di atas menyajikan analisis arah korelasi dan distribusi dampak dari fitur-fitur yang ada terhadap prediksi konversi. Sebagaimana

tervisualisasi, distribusi titik data (*Summary Plot*) mengindikasikan secara visual bahwa setiap penambahan nominal anggaran berkorelasi langsung dengan proyeksi peningkatan konversi.



Gambar 5. rafik SHAP Dependence Plot untuk

Variabel Pengeluaran Anggaran (*marketing_spend*)

Pada Gambar 4, *Dependence Plot* memetakan tren linier positif, yang menegaskan bahwa seiring dengan peningkatan investasi pemasaran (*marketing spend*), dampak terhadap metrik konversi (nilai SHAP) terus meningkat secara proporsional, tanpa menunjukkan gejala pelemahan utilitas marginal (*diminishing returns*). Di tingkat granular, analisis lebih lanjut melalui *Waterfall Plot* sukses menguraikan konstruksi probabilitas konversi pada sampel kampanye tunggal, mengkuantifikasi kontribusi aditif dari masing-masing fitur mulai dari nilai performa rata-rata (*base value*) hingga menembus kalkulasi akhir model.

4.2. Pembahasan

Penelitian ini dirancang untuk yang masih sering terjadi dalam praktik analisis pemasaran digital. Selama ini, pemodelan data pemasaran selalu terjebak dalam dilema menggunakan model

statistik linier tradisional yang mudah dipahami tetapi kurang akurat, atau menggunakan model algoritma canggih yang sangat akurat tetapi sulit dijelaskan alasannya. Kebaruan dari penelitian ini terletak pada integrasi algoritma CatBoost yang mampu memproses data pemasaran kategori dengan sangat baik bersama kerangka interpretasi SHAP. Penggabungan kedua metode ini terbukti memecahkan dilema tersebut, penelitian berhasil menghasilkan model dengan akurasi prediksi yang luar biasa tinggi (R^2 mencapai 99,12%) sekaligus menghasilkan penjelasan yang sangat transparan dan mudah dicerna oleh bahasa bisnis.

Berdasarkan hasil interpretasi komputasi yang direpresentasikan pada visualisasi SHAP, analisis ini membenarkan hipotesis awal bahwa kucuran anggaran pemasaran adalah penggerak utama yang paling menentukan sukses tidaknya sebuah kampanye. Terdapat temuan yang sangat menarik dari kacamata bisnis yang tidak ditemukan adanya gejala pelemahan hasil (*diminishing returns*). Kurva interaksi terus menunjukkan tren naik yang stabil hingga menyentuh batas anggaran pengamatan sebesar \$20.000 artinya penambahan anggaran dalam rentang nominal tersebut dipastikan akan terus memberikan peningkatan konversi yang sepadan tanpa mencapai titik jenuh penonton. Temuan ini memberikan dasar empiris yang sangat kuat bagi manajemen perusahaan untuk tidak ragu

menambah atau menaikkan skala anggaran (*scaling up*) pada kampanye yang di awal sudah menunjukkan kinerja positif. Selain itu, model analitik ini juga menyoroti pentingnya peran saluran pemasaran yang dipilih. Walaupun dampaknya tidak sebesar nominal anggaran, pemilihan media iklan sangat menentukan kestabilan performa kampanye. Penggunaan saluran dengan interaksi visual yang tinggi seperti Meta dan TikTok memiliki kecenderungan efek positif yang lebih konsisten dibandingkan saluran komunikasi yang kaku seperti Email. Manfaat praktis terbesar dari model ini dapat dilihat pada tingkat operasional audit kampanye. Fitur ini merevolusi cara para manajer pemasaran melakukan evaluasi. Dengan melihat penjabaran nilai dari satu kampanye tunggal secara detail, tim dapat membedah setiap dolar yang dikeluarkan dan melacak dampak langsungnya terhadap prediksi konversi. Hal ini menghilangkan unsur tebak-tebakan dan memberikan keyakinan penuh dalam merumuskan strategi alokasi dana untuk kuartal berikutnya.

5. Kesimpulan

Penelitian ini berhasil membuktikan keandalan kombinasi algoritma CatBoost dan instrumen SHAP dalam menciptakan sistem analisis pemasaran yang sangat presisi ($R^2 = 0,9912$) sekaligus transparan. Temuan empiris menegaskan bahwa alokasi anggaran merupakan

motor penggerak utama konversi tanpa menunjukkan gejala pelemahan efek (*diminishing returns*) hingga batas \$20.000, sementara pemilihan platform seperti Meta dan TikTok memberikan efek pengganda keberhasilan secara signifikan. Oleh karena itu, jajaran manajemen direkomendasikan untuk mengadopsi strategi ekspansi anggaran secara agresif pada saluran digital yang terbukti efektif. Guna memperkaya landasan analitik di masa mendatang, model ini dapat dikembangkan lebih lanjut dengan menyertakan indikator Biaya Per Klik (CPC) dan Biaya Akuisisi Pelanggan (CAC), serta diintegrasikan ke dalam wujud aplikasi dasbor interaktif untuk memfasilitasi simulasi perencanaan anggaran secara real-time.

6. Daftar Pustaka

- [1] N. Laziva and M. Q. Atieq, "Studi Literatur Digitalisasi UMKM (Usaha Mikro, Kecil, dan Menengah) di Era Masyarakat 5.0: Strategi dan Faktor," *Magister Manaj. Fak. Ekon. dan Bisnis Univ. Tanjungpura*, vol. 7, 2024.
- [2] S. Kemp, "Digital 2025: global advertising trends," *datareportal*.
- [3] A. K. Virgiawan, E. D. Absharina, and F. Fenando, "peran big data dalam meningkatkan daya saing bisnis di era digital," *Simtek J. Sist. Inf. dan Tek. Komput.*, vol. 10, no. 1, 2025, doi: 10.51876/simtek.v10i1.1378.
- [4] J. A. Apaza-Caceres, L. A. V. Paja, A. P. Huachaca, and D. A. P. Choque, "Machine learning applications in digital advertising performance optimization: A systematic literature review," *Int. J. Data Netw. Sci.*, vol. 10, no. 1, 2026, doi:

- 10.5267/j.ijdns.2025.9.020.
- [5] I. Wayan, Budiman, Casriyanti, I. Megawati, and E. Al, *Buku Ajar Manajemen Pemasaran*, vol. 11, no. 1. 2023.
- [6] R. Shibiti, "Measuring the Return on Investment (ROI) through an integrated framework: A case of academic staff training and development interventions in South Africa," *Int. J. Res. Bus. Soc. Sci.* (2147- 4478), vol. 14, no. 9, 2026, doi: 10.20525/ijrbs.v14i9.4621.
- [7] T. Patrocínio, M. Madaleno, and M. C. Nogueira, "Does the Way Variables Are Calculated Change the Conclusions to Be Drawn? A Study Applied to the Ratio ROI (Return on Investment)," *J. Risk Financ. Manag.*, vol. 17, no. 7, 2024, doi: 10.3390/jrfm17070266.
- [8] R. Almestarihi, A. Y. A. B. Ahmad, R. H. Frangieh, I. A. Abualsondos, K. K. Nser, and A. Ziani, "Measuring the ROI of paid advertising campaigns in digital marketing and its effect on business profitability," *Uncertain Supply Chain Manag.*, vol. 12, no. 2, 2024, doi: 10.5267/j.uscm.2023.11.009.
- [9] L. Sekar, D. Hartanti, and A. A. Sari, "Prototype Sistem Rekomendasi Pemilihan Produk Furniture dengan Pemodelan Content-Based Filtering," *J. Ris. dan Apl. Mhs. Inform.*, vol. 6, no. 01, 2025, doi: 10.30998/jrami.v6i01.11302.
- [10] S. Suliati, S. Achmadi, and D. Rudhistiar, "penerapan sistem pakar untuk deteksi dini mental illness dengan menggunakan metode forward chaining dan certainty factor berbasis website," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 6, no. 2, 2023, doi: 10.36040/jati.v6i2.5464.
- [11] I. Jaya and S. Hanadwiputra, "Penerapan Algoritma Apriori Dalam Menentukan Strategi Persediaan Obat Terhadap Penjualan Obat Pada Apotek Karang Asih," *Qual. Syst. Dev.*, vol. 11 No 1, no. 2021-02-08, 2021.
- [12] N. T. Mukomberanwa *et al.*, "Wildlife-vehicle collisions in the Mid-Zambezi Valley: Driver perceptions and species-specific risk patterns," *Sci. African*, vol. 31, 2026, doi: 10.1016/j.sciaf.2025.e03137.
- [13] J. Wang and S. Zhou, "Particle swarm optimization-XGBoost-based modeling of radio-frequency power amplifier under different temperatures," *Int. J. Numer. Model. Electron. Networks, Devices Fields*, vol. 37, no. 2, 2024, doi: 10.1002/jnm.3168.
- [14] M. R. Wijaya and G. S. Wibowo, "Customer Segmentation berdasarkan Usia, Jumlah Kredit dan Lama Kredit Nasabah di Bank XYZ menggunakan Model K-Means Clustering," *Pros. Semin. Nas. Univ. Ma Chung*, vol. 1, 2021, doi: 10.33479/snumc.v1i.228.
- [15] Y. Yang, "A Study on Bank Marketing Prediction Based on XGBoost," *Adv. Econ. Manag. Polit. Sci.*, vol. 185, no. 1, 2025, doi: 10.54254/2754-1169/2025.lh23949.
- [16] J. Cui, H. Jiang, and Z. Xu, "Digital marketing program design based on abnormal consumer behavior data classification and improved homomorphic encryption algorithm," *PeerJ Comput. Sci.*, vol. 9, 2023, doi: 10.7717/peerj-cs.1690.
- [17] Q. Yu, "Enhancing Bank Term Deposit Predictions: A Machine Learning Approach with CatBoost and SHAP," *Appl. Comput. Eng.*, vol. 120, no. 1, 2025, doi: 10.54254/2755-2721/2025.19485.
- [18] F. Faiaz, T. K. Tapu, A. Nawer, and S. R. Payel, "Transparent Loan Approval Prediction: A Machine Learning Approach with SHAP-based Interpretability," in *2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025*, 2025, doi: 10.1109/ECCE64574.2025.11014000.
- [19] J. C. Okoye *et al.*, "Integrating Explainable AI and Customer Behaviour Modelling for Strategic Insights in E-Commerce and Business Intelligence Cite as," *Int. J. Strateg. Manag. Bus. Policy*, vol. 7, no. 2, pp. 8–15, 2025, doi:

- 10.5281/zenodo.17115504.
- [20] J. Liu and C. Liu, "Prediction of Online Shoppers' Purchasing Intention Based on SHAP-IGWO-EM Method," *Int. J. Swarm Intell. Res.*, vol. 17, no. 1, 2026, doi: 10.4018/ijisir.397667.
- [21] X. Cheng *et al.*, "CatBoost-based prediction of suspended sediment concentration in the Pearl River estuary: Driving mechanisms unraveled via SHAP analysis," *J. Sea Res.*, vol. 210, 2026, doi: 10.1016/j.seares.2026.102668.
- [22] O. K. Sari and A. Cherid, "Aplikasi Berbasis Web Menggunakan Apriori Untuk Rekomendasi Bundling Produk Sembako," *Bit (Fakultas Teknol. Inf. Univ. Budi Luhur)*, vol. 20, no. 2, 2023, doi: 10.36080/bit.v20i2.2444.
- [23] A. Wadanur and A. A. Sari, "Implementasi Algoritma Apriori dan FP-Growth pada Penjualan Spareparts," *Edumatic J. Pendidik. Inform.*, vol. 6, no. 1, pp. 107–115, 2022, doi: 10.29408/edumatic.v6i1.5470.
- [24] A. A. Sari, I. T. Saputra, and A. D. Prakoso, "Optimalisasi Proses Digitalisasi UMKM melalui Aplikasi Marketplace berbasis Design Thinking," vol. 8, no. 2, pp. 535–544, 2024, doi: 10.29408/edumatic.v8i2.27702.
- [25] S. Ben Jabeur, S. Meftah-Wali, and J. L. Viviani, "Forecasting gold price with the XGBoost algorithm and SHAP interaction values," *Ann. Oper. Res.*, vol. 334, no. 1–3, 2024, doi: 10.1007/s10479-021-04187-w.
- [26] Hasim Azari, Dwi Hartanti, and Aprilisa Arum Sari, "Pengelompokan Produksi Padi dan Beras Provinsi Jawa Timur dengan Metode Agglomerative Hierarchical Clustering," *Infotek J. Inform. dan Teknol.*, vol. 7, no. 2, 2024, doi: 10.29408/jit.v7i2.26016.
- [27] J. Zhang and Z. Zhao, "Corporate ESG rating prediction based on XGBoost-SHAP interpretable machine learning model," *Expert Syst. Appl.*, vol. 295, 2026, doi: 10.1016/j.eswa.2025.128809.
- [28] W. Lu, "Research on the Optimization of Bank Precision Marketing Strategies Based on Catboost," *Adv. Econ. Manag. Polit. Sci.*, vol. 200, no. 1, 2025, doi: 10.54254/2754-1169/2025.lh25226.
- [29] X. Ke *et al.*, "From prediction to regionalization: Enhancing flash flood susceptibility mapping using machine learning and GeoDetector," *Geosci. Front.*, vol. 17, no. 1, 2026, doi: 10.1016/j.gsf.2025.102213.
- [30] T. C. Vo, T. Q. Nguyen, and V. L. Tran, "Predicting and optimizing the concrete compressive strength using an explainable boosting machine learning model," *Asian J. Civ. Eng.*, vol. 25, no. 2, 2024, doi: 10.1007/s42107-023-00848-2.
- [31] W. E. E. Ebri, B. A. E. Bassey, and E. J. K. Enyam, "A Comparative Analysis Of Traditional Vs Digital Advertising Impacts On Consumers' Purchase Behaviour," *J. Adm. Manag. Sci.*, vol. 1, no. 1, 2026, doi: 10.64415/jams.v1i1.32.
- [32] S. P. S. Patiro, D. Aryani, E. Rekarti, and M. Wibowo, "Peningkatan pengetahuan marketplace pada pelaku usaha mikro kecil dan menengah (UMKM) dengan website education marketplace," *J. Inov. Has. Pengabd. Masy.*, vol. 6, no. 1, 2023, doi: 10.33474/jipemas.v6i1.19282.
- [33] I Wayan Willy Mustika and Salsa Bila Jihan Maulidah, "Analisis Penggunaan Media Sosial Sebagai Sarana Pemasaran pada Usaha Kecil Menengah," *J. Ris. Manaj. Komun.*, vol. 3, no. 1, 2023, doi: 10.29313/jrmk.v3i1.1716..